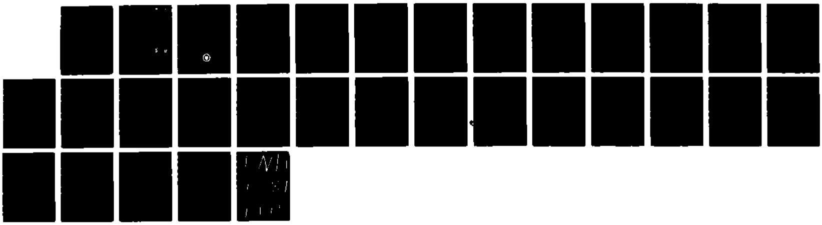
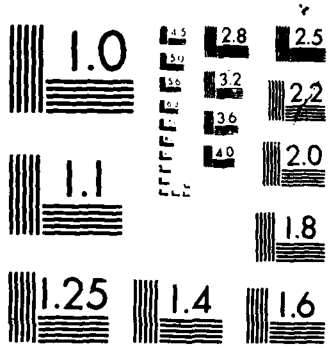
AD-A185 501 ON DETECTION OF CHANGE POINTS USING MEAN VECTORS(U) 1/1
PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS
P R KRISHNAIAH ET AL. DEC 86 TR-86-41 AFOSR-TR-87-1020
UNCLASSIFIED F49620-85-C-0008 F/G 12/9 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD-A185 581

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 87-1020	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) "On detection of change points using mean vectors"		5. TYPE OF REPORT & PERIOD COVERED ~~technical~~ Final - December 1986
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) P.R. Krishnaiah, B.Q. Miao and L.C. Zhao		8. CONTRACT OR GRANT NUMBER(s) F49620-85-C-0008
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis, 515 Thackeray Hall, University of Pittsburgh, Pittsburgh, PA 15260		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304 A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research, Department of the Air Force, Bolling Air Force Base, DC 20332		12. REPORT DATE December 1986
		13. NUMBER OF PAGES 27
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Same as 11		15. SECURITY CLASS. (of this report) unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Change points, Consistency, Edge detection, Information theoretic criterion, Model selection, Quality control.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper, the authors consider the problem of change points within the framework of model selection procedures using information theoretic criteria. The authors proposed procedures for estimation of the locations of change points and the number of change points. The strong consistency of these procedures is also established. Also, the problem of change points is discussed within the framework of the simultaneous test procedures.

unclassified

AFOSR-TR- 87-1020

ON DETECTION OF CHANGE POINTS
USING MEAN VECTORS*

by

P.R. Krishnaiah, B.Q. Miao
and L.C. Zhao
Center for Multivariate Analysis
University of Pittsburgh

Center for Multivariate Analysis

University of Pittsburgh

87 9 24 15

ON DETECTION OF CHANGE POINTS
USING MEAN VECTORS*

by

P.R. Krishnaiah, B.Q. Miao
and L.C. Zhao
Center for Multivariate Analysis
University of Pittsburgh

Technical Report No. 86-41

December 1986

Center for Multivariate Analysis
University of Pittsburgh
515 Thackeray Hall
Pittsburgh, PA 15260

*This work is supported by Contract F49620-85-C-0008 of the Air force Office of Scientific Research.

ABSTRACT

In this paper, the authors consider the problem of change points within the framework of model selection procedures using information theoretic criteria. The authors proposed procedures for estimation of the locations of change points and the number of change points. The strong consistency of these proceudres is also established. Also, the problem of change points is discussed within the framework of the simultaneous test procedures.

Keywords and phrases: Change points, Consistency, Edge detection, Information theoretic criterion, Model selection, Quality control.

1. INTRODUCTION

Problems of detection of change points arise in many areas. For example, in continuous production process, it is of interest to find out the point at which deterioration in the quality of the product starts. Change point problems arise (see Mazumdar, Sinha and Li (1985)) in the area of edge detection. For discussions on some other applications, the reader is referred to Page (1957) and Holbert and Broemling (1977).

When the underlying distribution is normal, the change point problem is equivalent to detection of change in mean and/or variance. Chernoff and Zacks (1964), Page (1955), Hinkley (1970) and several other workers investigated the problem of change point in the mean. Sen and Srivastava (1973) and Srivastava and Worsley (1986) investigated the problem of detection of the change in the mean vectors. In the above papers, change point problems were studied within the framework of tests of hypotheses.

The object of this paper is to study change points in the mean vectors within the framework of model selection procedures using information theoretic criteria. In Section 2, we give some preliminaries which are needed in the sequel. Section 3 is devoted to the problem of estimation of the locations of change points when the number of change points is known. The strong consistency of the procedure is also established. In Section 4, we assume that the number of change points is unknown and propose procedures for the estimation of the number and locations of change points. The strong consistency of these procedures is also established. Finally, in Section 5, we discussed the change point problem within the framework of simultaneous test procedures.

2. PRELIMINARIES

Let $\underset{\sim}{x}(t)$ be an independent p-dimensional process on [0,1] such that

$$\underset{\sim}{x}(t) = \underset{\sim}{\mu}(t) + \underset{\sim}{v}(t), \quad 0 \leq t \leq 1 \tag{2.1}$$

where $\underset{\sim}{\mu}(t)$ is a non-random p-dimensional left-continuous step function and $\underset{\sim}{v}(t)$ is an independent p-variate normal process with mean vector $\underset{\sim}{0}$ and covariance matrix Σ. Denote the jump points of $\underset{\sim}{\mu}(t)$ by $t_1,\ldots,t_q$;$(0 < t_1 < t_2 < \ldots < t_q < 1)$. Then $t_1,\ldots,t_q$ are called the change points of the process $\underset{\sim}{x}(t)$. We assume that N samples are drawn from the process $\underset{\sim}{x}(t)$ in equal space, say $\underset{\sim}{x}(\frac{1}{N}),\ldots,\underset{\sim}{x}(\frac{N}{N})$. Then, it is of interest to estimate the location of $(t_1,\ldots,t_q)$ from observations $\underset{\sim}{x}(\frac{1}{N}),\ldots,\underset{\sim}{x}(\frac{N}{N})$. Here it is emphasized that for different N, the series $\{\underset{\sim}{x}(\frac{1}{N}),\ldots,\underset{\sim}{x}(\frac{N}{N})\}$ is different. Sometimes, the number q of change points is unknown and in this case we have to estimate q. In this paper, it is assumed a priori that $q \leq L$, a constant.

We now define some concepts and introduce some notations which are useful in the sequel. From the integer interval [0,N], we can pick ℓ integers from [1,N-1] to divide [0,N] into $\ell + 1$ sections. Let $k_1 < \ldots < k_\ell$ $(0<k_i<N)$ denote the ℓ integers picked from [1,N-1]. Then, we call $(k_1,\ldots,k_\ell)$ a partition of the interval [0,N], and k_i, $1 \leq i \leq \ell$, a cut off point of this partition. Sometimes, we denote $(k_1,\ldots,k_\ell)$ by π_ℓ or π simply. A partition $(z_1,\ldots,z_i)$ is called a refinement partition of $(k_1,\ldots,k_\ell)$ if the set $\{k_1,k_2,\ldots,k_\ell\}$ is a proper subset of $\{z_1,\ldots,z_i\}$. This fact is written as $(z_1,\ldots,z_i)$ $(k_1,\ldots,k_\ell)$. If we define $\pi_i = (z_1,\ldots,z_i)$, $\pi_\ell = (k_1,\ldots,k_\ell)$, the above relation is written as $\pi_i \prec \pi_\ell$.

Define $\underset{\sim}{x}_i^{(N)} = \underset{\sim}{x}(\frac{i}{N})$, $1 \leq i \leq N$. It is obvious that $\underset{\sim}{x}_i^{(N)}$ is dependent on the sample number N. Throughout this paper, we will use $\underset{\sim}{x}_i$ instead of $\underset{\sim}{x}_i^{(N)}$, but we must keep in mind that $\underset{\sim}{x}_i$ is dependent on N.

If $\underset{\sim}{x}$ is a vector, then $\underset{\sim}{x}'$ denotes the transpose of $\underset{\sim}{x}$. Also, the maximum and minimum eigenvalues of the matrix M are denoted by $\lambda_{max}(M)$ and $\lambda_{min}(M)$ respectively. Throughout this paper, C_1, C_2, C_3 etc. denote some positive constants which can assume different values at each of their appearances, and the Eculidean norm of M is denoted by $||M||$. Now we shall prove some lemmas. At first, we cite a lemma by Fuk and Nageav (1971).

<u>Lemma 2.1</u>. Let $x_1, \ldots, x_n$ be independent random variables with $Ex_i = 0$, and $E|x_i|^t < \infty$, $i = 1,2,\ldots,n$. For $t \geq 2$, write $S_n = \sum_{i=1}^{n} x_i$, $B_n^2 = \sum_{i=1}^{n} Var(x_i)$, $A_{t,n} = \sum_{i=1}^{n} E|x_i|^t$. Then for $x > 0$,

$$P(S_n \geq x) \leq C_t^{(1)} A_{t,n} x^{-t} + \exp\{-C_t^{(2)} x^2/B_n^2\}$$

where

$$C_t^{(1)} = (1+\frac{2}{t})^t \quad \text{and} \quad C_t^{(2)} = 2(t+2)^{-2}e^{-t}.$$

In the original paper, $C_t^{(2)} = 2(t+2)^{-1}e^{-t}$; this is a printing error.

<u>Lemma 2.2</u>. Let $\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_N$ be iid p×1 normal vectors with mean vector $\underset{\sim}{0}$ and positive definite covariance matrix Σ. Now, let

$$A_N = \frac{1}{N} \sum_{i=1}^{N} (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{0N})(\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{0N})' \tag{2.2}$$

where

$$\bar{\underset{\sim}{x}}_{ij} = \frac{1}{j-i} \sum_{k=i+1}^{j} \underset{\sim}{x}_k . \tag{2.3}$$

Then,

$$P\{\lambda_{max}(A_N) \geq \frac{3}{2}\lambda_{max}(\Sigma)\} < CN^{-(L+2)} \tag{2.4}$$

$$P\{\lambda_{min}(A_N) \leq \frac{1}{2}\lambda_{min}(\Sigma)\} < CN^{-(L+2)} \tag{2.5}$$

Proof. For the sake of simplicity, write $\bar{x}_{0N} = \bar{x}$, $\bar{x}_{0N_1} = \bar{x}_1$, $\bar{x}_{N_1N} = \bar{x}_2$. Then

$$A_N = \frac{1}{N}\sum_{i=1}^{N} x_i x_i' - \bar{x}\bar{x}'. \tag{2.6}$$

Let

$$x_i = (x_{i1},\ldots,x_{ip})', \qquad 1 \leq i \leq N$$

$$y_{jk} = \frac{1}{N}\sum_{i=1}^{N} x_{ij}x_{ik} \qquad 1 \leq j,k \leq p. \tag{2.7}$$

It is easy to see that y_{jk} is the average of N iid random variables. Its fourth moment is finite. Let $Ex_{1j}x_{1k} = \mu_{jk}$ and put $t = 4$ in the Lemma 2.1. Then for any $\varepsilon > 0$,

$$P\left(\frac{1}{N}\left|\sum_{i=1}^{N}(x_{ij}x_{ik}-\mu_{jk})\right| \geq \frac{\varepsilon}{2p^2}\right) < C_1N^{-(L+2)}. \tag{2.8}$$

and

$$P(||\bar{x}\bar{x}'|| \geq \varepsilon/2) < C_1N^{-(L+2)}.$$

But

$$\Sigma = (\mu_{jk}):p\times p \text{ and } \frac{1}{N}\sum_{i=1}^{N} x_i x_i' = \left(\frac{1}{N}\sum_{i=1}^{N} x_{ij}x_{ik}\right):p\times p.$$

It follows

$$P\{||A_N - \Sigma|| \geq \varepsilon\} < C_2N^{-(L+2)} \tag{2.9}$$

Let $\lambda_1(A) \geq \cdots \geq \lambda_p(A)$ denote the eigenvalues of a p×p matrix A in the sequel. Then a well-known inequality about symmetric matrix, we have

$$\sum_{i=1}^{p} (\lambda_i(A_N)-\lambda_i(\Sigma))^2 \leq \operatorname{tr}(A_N-\Sigma)^2 \tag{2.10}$$

The relations of (2.4) and (2.5) follow immediately from (2.9) and (2.10).

Lemma 2.3. Let $\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_N$ be iid p×1 normal vectors, (again take notice of that $\underset{\sim}{x}_i$, $1 \leq i \leq N$, is dependent on N), $\underset{\sim}{x}_1 \sim N_p(\underset{\sim}{0},\Sigma)$, $\Sigma > 0$ and is unknown. Write

$$A_N = \frac{1}{N} \sum_{i=1}^{N} (\underset{\sim}{x}_i-\bar{\underset{\sim}{x}}_{0N})(\underset{\sim}{x}_i-\bar{\underset{\sim}{x}}_{0N})',$$

$$B_N = \frac{1}{N} \sum_{i=1}^{N_1}(\underset{\sim}{x}_i-\bar{\underset{\sim}{x}}_{0N_1})(\underset{\sim}{x}_i-\bar{\underset{\sim}{x}}_{0N_1})' + \frac{1}{N} \sum_{i=N_1+1}^{N} (\underset{\sim}{x}_i-\bar{\underset{\sim}{x}}_{N_1N})(\underset{\sim}{x}_i-\bar{\underset{\sim}{x}}_{N_1N})' \tag{2.11}$$

where

$$\bar{\underset{\sim}{x}}_{ij} = \frac{1}{j-i} \sum_{k=i+1}^{j} \underset{\sim}{x}_k.$$

Then we have that A_N-B_N is non-negative definite, and for k_1, k_2, $0 \leq k_1 < k_2 \leq N$, the following estimates hold.

$$P\{\lambda_{max}(\frac{k_2-k_1}{N} \bar{\underset{\sim}{x}}_{k_1k_2}\bar{\underset{\sim}{x}}'_{k_1k_2}) \geq C \frac{\log N}{N} \} < C_1N^{-(L+2)}, \tag{2.12}$$

$$P\{ \frac{N}{2} (\log|A_N|-\log|B_N|) \geq C \log N\} < C_1 N^{-(L+2)} \tag{2.13}$$

Proof. Without loss of generality, we prove (2.12) for $k_1 = 0$, $k_2 = N$ only. Let $\underset{\sim}{y}^{(N)} = \sqrt{N}\, \Sigma^{-\frac{1}{2}}\bar{\underset{\sim}{x}}_{0N}$. Then $\underset{\sim}{y}^{(N)} \sim N_p(\underset{\sim}{0},I_p)$. Write $\underset{\sim}{y}^{(N)} = (y_1^{(N)},\ldots,y_p^{(N)})'$, and $\underset{\sim}{y}^{(N)}\underset{\sim}{y}^{(N)\prime} = (y_{ij}^{(N)}):p\times p$, where $y_{ij}^{(N)} = y_i^{(N)}y_j^{(N)}$. Since $y_i^{(N)} \sim N(0,1)$, for all i, $1 \le i \le p$ it is not difficult to see that

$$P(|y_i^{(N)}| > C\log^{\frac{1}{2}}N) \le \frac{1}{\sqrt{2\pi}\ C\log^{\frac{1}{2}}N}\, e^{-\frac{C}{2}\log N} < C_1N^{-(L+2)}$$

for some constant C, and $N \ge 3$.

From the fact that $2|y_iy_j| \le y_i^2 + y_j^2$, it follows immediately that

$$P\{\|\underset{\sim}{y}^{(N)}\underset{\sim}{y}^{(N)\prime}\| \ge C\log N\} < C_1N^{-(L+2)}.$$

So

$$P\{\|N\bar{\underset{\sim}{x}}_{ON}\bar{\underset{\sim}{x}}_{ON}'\| \ge C\log N\} < C_1N^{-(L+2)},$$

which implies (2.12).

Here, we note that

$$A_N = \frac{1}{N}\sum_{i=1}^{N}(\underset{\sim}{x}_i-\bar{\underset{\sim}{x}}_{ON})(\underset{\sim}{x}_i-\bar{\underset{\sim}{x}}_{ON})'$$

$$= B_N + \frac{N_1(N-N_1)}{N^2}(\bar{\underset{\sim}{x}}_{ON_1}-\bar{\underset{\sim}{x}}_{N_1N})(\bar{\underset{\sim}{x}}_{ON_1}-\bar{\underset{\sim}{x}}_{N_1N})'$$

and for any vectors $\underset{\sim}{x}$ and $\underset{\sim}{y}$

$$\tfrac{1}{2}\,\underset{\sim}{x}\underset{\sim}{x}' - \underset{\sim}{y}\underset{\sim}{y}' \le (\underset{\sim}{x}-\underset{\sim}{y})(\underset{\sim}{x}-\underset{\sim}{y})' \le 2(\underset{\sim}{x}\underset{\sim}{x}' + \underset{\sim}{y}\underset{\sim}{y}') \qquad (2.14)$$

we have

$$0 \le A_N - B_N \le \frac{2N_1N_2}{N^2}(\bar{\underset{\sim}{x}}_{ON_1}\bar{\underset{\sim}{x}}_{ON_1}' + \bar{\underset{\sim}{x}}_{N_1N}\bar{\underset{\sim}{x}}_{N_1N}').$$

From (2.12), we get

$$P\{\lambda_{max}(A_N - B_N) \geq C \frac{\log N}{N}\} < C_1 N^{-(L+2)} \tag{2.15}$$

It is well-known that for any symmetric matrix D_2 and $D_1 > 0$,

$$\frac{\lambda_{max}(D_2)}{\lambda_{max}(D_1)} \leq \lambda_{max}(D_1^{-\frac{1}{2}} D_2 D_1^{-\frac{1}{2}}) \leq \frac{\lambda_{max}(D_2)}{\lambda_{min}(D_1)} . \tag{2.16}$$

By Lemma 2.2, we have

$$P\{\lambda_{min}(B_N) \leq \frac{1}{2} \lambda_{min}(\Sigma)\} < C_1 N^{-(L+2)}. \tag{2.17}$$

From (2.15) - (2.17), it follows that

$$P\{\delta_i \geq C \log N/N\} < C_1 N^{-(L+2)}$$

where $\delta_i \hat{=} \lambda_i(B_N^{-\frac{1}{2}}(A_N - B_N)B_N^{-\frac{1}{2}})$.

From

$$\frac{N}{2}(\log|A_N| - \log|B_N|) = \frac{N}{2} \log|(I + B_N^{-\frac{1}{2}}(A_N - B_N)B_N^{-\frac{1}{2}})|$$

$$= \frac{N}{2} \log(\prod_{i=1}^{p} (1+\delta_i)) \leq \frac{N}{2} \sum_{i=1}^{p} \delta_i$$

(2.13) follows.

Remark 2.1. Let $\pi_\ell = (k_1, \ldots, k_\ell)$ be a partition of $[0,N]$, where ℓ is less than a certain constant L. Set

$$A_{\pi_\ell}(N) = \frac{1}{N} \sum_{j=0}^{\ell} \sum_{i=k_j+1}^{k_{j+1}} (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{k_j k_{j+1}})(\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{k_j k_{j+1}})', \tag{2.18}$$

where $k_o = 0$, $k_{\ell+1} = N$. Then

$$\begin{aligned} &P\{\lambda_{max}(A_{\pi_\ell}(N)) \geq \frac{3}{2} \lambda_{max}(\Sigma)\} < C_1 N^{-(L+2)}, \\ &P\{\lambda_{min}(A_{\pi_\ell}(N)) \leq \frac{1}{2} \lambda_{min}(\Sigma)\} < C_1 N^{-(L+2)}. \end{aligned} \tag{2.19}$$

To see that (2.19) hold true, we observe that

$$A_N - A_{\pi_\ell}(N) = \frac{1}{N}\sum_{j=0}^{\ell}(k_{j+1}-k_j)(\bar{x}_{k_jk_{j+1}}-\bar{x}_{0N})(\bar{x}_{k_jk_{j+1}}-\bar{x}_{0N})'$$

$$\leq 2\sum_{j=0}^{\ell}\frac{(k_{j+1}-k_j)}{N}\bar{x}_{k_jk_{j+1}}\bar{x}'_{k_jk_{j+1}} + \frac{2}{N}\sum_{j=0}^{\ell}(k_{j+1}-k_j)\bar{x}_{0N}\bar{x}'_{0N}$$

$$= 2\sum_{j=0}^{\ell}\frac{k_{j+1}-k_j}{N}\bar{x}_{k_jk_{j+1}}\bar{x}'_{k_jk_{j+1}} + 2\bar{x}_{0N}\bar{x}'_{0N}.$$

From (2.12), we have

$$P\{\lambda_{\max}(A_N-A_{\pi_\ell}(N)) \geq C\frac{\log N}{N}\} < C_1N^{-(L+2)} \tag{2.20}$$

Write $\lambda_i(A_N-A_{\pi_\ell}(N)) = \lambda_i$ simply and write $A_N - A_{\pi_\ell}(N) = (\alpha_{ij}): p\times p$

By means of inequalitites

$$\lambda_1 \geq \max_{1\leq i\leq p}\alpha_{ii}$$

$$\lambda_1\lambda_2 \geq \max_{1\leq i<j\leq p}(\alpha_{ii}\alpha_{jj}-\alpha_{ij}^2),$$

we have for any $\varepsilon > 0$,

$$P\{||A_N - A_{\pi_\ell}(N)|| \geq \varepsilon\} < C_1N^{-(L+2)} \tag{2.21}$$

Combining (2.9) and (2.20), we have

$$P\{||A_{\pi_\ell}(N) - \Sigma|| \geq \varepsilon\} < C_1N^{-(L+2)} \tag{2.22}$$

Thus, we obtain (2.19).

Remark 2.2. Let $x_1,\ldots,x_N$ be independent $p\times 1$ normal vectors with common covariance matrix $\Sigma > 0$. Assume that

$$Ex_1 = \ldots = Ex_{k_1} = \mu_1,\quad Ex_{k_1+1} = \ldots = Ex_{k_2} = \mu_2,\ \ldots$$

$$Ex_{k_{\ell-1}+1} = \ldots = E(x_{k_\ell}) = \mu_\ell,\ Ex_{k_\ell+1} = \ldots = Ex_N = \mu_{\ell+1}.$$

and

$$B_{\pi_\ell}(N) = \frac{1}{N} \sum_{j=0}^{\ell} \sum_{i=k_j+1}^{k_{j+1}} (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{k_j k_{j+1}})(\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{k_j k_{j+1}})'$$

Then for $\pi_\ell = (k_1,\ldots,k_\ell)$, we have for $\varepsilon > 0$

$$P\{||B_{\pi_\ell}(N)-\Sigma|| \geq \varepsilon\} < C_1 N^{-(L+2)} \tag{2.23}$$

The proof is very easy if we notice that $\underset{\sim}{x}_1-\underset{\sim}{\mu}_1,\ldots,\underset{\sim}{x}_{k_1}-\underset{\sim}{\mu}_1$, $\underset{\sim}{x}_{k_1+1}-\underset{\sim}{\mu}_2,\ldots,$ $\underset{\sim}{x}_{k_\ell+1}-\underset{\sim}{\mu}_{\ell+1}$ are iid p×1 normal vectors so that we can apply Remark 2.1.

<u>Lemma 2.4</u>. Let $\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_N$ be independent p×1 normal vectors with the same covariance matrix $\Sigma > 0$, and $\pi_\ell = (k_1,\ldots,k_\ell)$ be a partition of $[0,N]$. Assume

$$E\underset{\sim}{x}_i = \mu_j, \quad k_{j-1} < i \leq k_j, \quad \underset{\sim}{\mu}_j \neq \underset{\sim}{\mu}_{j+1} \tag{2.24}$$

where $j = 1,2,\ldots\ell+1$, and $k_o = 0$, $k_{\ell+1} = N$. Set

$$A_N(k_j,N_1,N_2) = \frac{1}{N} \sum_{i=k_j-N_1+1}^{k_j+N_2} (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{k_j-N_1,k_j+N_2})(\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{k_j-N_1,k_j+N_2})' \tag{2.25}$$

$$B_N(k_j,N_1,N_2) = \frac{1}{N} \sum_{i=k_j-N_1+1}^{k_j} (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{k_j-N_1,k_j})(\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{k_j-N_1,k_j})'$$

$$+ \frac{1}{N} \sum_{i=k_j+1}^{k_j+N_2} (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{k_j,k_j+N_2})(\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{k_j,k_j+N_2})' \tag{2.26}$$

where $N_1 \leq k_j - k_{j-1}$, and $N_2 \leq k_{j+1} - k_j$. Set $\pi_{\ell+1} = (k_1,\ldots,k_{j-1},k_j-N_1,$ $k_j+N_2,\ k_{j+1},\ldots,k_\ell)$ and $\pi_{\ell+2} = (k_1,\ldots,k_{j-1},k_j-N_1,k_j,k_j+N_2,k_{j+1},\ldots,k_\ell)$. Suppose that for some positive constant β,

$$N_1 \geq \beta D_N, \quad N_2 \geq \beta D_N \tag{2.27}$$

where D_N satisfies the following conditions:

$$\text{(i)} \quad \lim_{N\to\infty} D_N/\log N = \infty, \tag{2.28}$$

$$\text{(ii)} \quad \lim_{N\to\infty} D_N/N = 0. \tag{2.29}$$

Then there exist constants C_1 and C_2 such that

$$P\{\lambda_{max}(A_N(k_j,N_1,N_2)-B_N(k_j,N_1,N_2)) < C_1 \frac{D_N}{N}\} < C_3 N^{-(L+2)}, \tag{2.30}$$

and

$$P\{\frac{N}{2}(\log|A_{\pi_{\ell+1}}|-\log|A_{\pi_{\ell+2}}|) < C_2 D_N\} < C_3 N^{-(L+2)}, \tag{2.31}$$

where $A_{\pi_{\ell+1}}$, $A_{\pi_{\ell+2}}$ are defined by (2.18).

Proof. By calculation and the inequality (2.14), we have

$$\begin{aligned}
A_{\pi_{\ell+1}} - A_{\pi_{\ell+2}} &= A_N(k_j,N_1,N_2) - B_N(k_j,N_1,N_2) \\
&= \frac{N_1N_2}{N(N_1+N_2)} (\bar{\underset{\sim}{x}}_{k_j-N_1,k_j} - \bar{\underset{\sim}{x}}_{k_j,k_j+N_2})(\bar{\underset{\sim}{x}}_{k_j-N_1,k_j} - \bar{\underset{\sim}{x}}_{k_j,k_j+N_2})' \\
&\geq \frac{N_1N_2}{N(N_1+N_2)} (\tfrac{1}{2}(\underset{\sim}{\mu}_{j+1}-\underset{\sim}{\mu}_j)(\underset{\sim}{\mu}_{j+1}-\underset{\sim}{\mu}_j)' - 2(\bar{\underset{\sim}{x}}_{k_j-M_1,k_j}-\underset{\sim}{\mu}_j)(\bar{\underset{\sim}{x}}_{k_j-N_1,k_j}-\underset{\sim}{\mu}_j)' \\
&\quad - 2(\bar{\underset{\sim}{x}}_{k_j,k_j+N_2}-\underset{\sim}{\mu}_{j+1})(\bar{\underset{\sim}{x}}_{k_j,k_j+N_2}-\underset{\sim}{\mu}_{j+1})') \\
&\hat{=} (J_1-J_2-J_3).
\end{aligned} \tag{2.32}$$

From Lemma 2.3, Eq:(2.29) and $\lambda_{max}(J_1) \geq \frac{1}{p}\,\mathrm{tr}(J_1)$, we get

$$P\{\lambda_{max}(J_2+J_3) \geq C \frac{\log N}{N}\} < C_1 N^{-(L+2)} \tag{2.33}$$

By

$$\lambda_{max}(J_1) \geq \frac{1}{2p} \frac{N_1N_2}{N(N_1+N_2)} \sum_{j=1}^{p} (\mu_{2j}-\mu_{1j})^2 \geq 2c_1 D_N/N, \tag{2.34}$$

where

$$c_1 = \frac{\beta}{8p} \sum_{j=1}^{p} (\mu_{2j}-\mu_{1j})^2 = \frac{\beta}{8p} ||\underset{\sim}{\mu}_{j+1}-\underset{\sim}{\mu}_j||^2 .$$

Combine (2.33) and (2.34), the inequality (2.30) follows. By (2.19),

$$P\{\lambda_{max}(A_{\pi_{\ell+2}}) > \frac{3}{2}\lambda_{max}(\Sigma)\} < C_1 N^{-(L+2)}, \tag{2.35}$$

$$P\{\lambda_{min}(A_{\pi_{\ell+2}}) < \frac{1}{2}\lambda_{min}(\Sigma)\} < C_1 N^{-(L+2)} .$$

By (2.14), (2.33) and expression of $A_{\pi_{\ell+1}} - A_{\pi_{\ell+2}}$,

$$P\{\lambda_{max}(A_{\pi_{\ell+1}}-A_{\pi_{\ell+2}}) \geq 2||\underset{\sim}{\mu}_{j+1}-\underset{\sim}{\mu}_j||^2\} < C_1 N^{-(L+2)} \tag{2.36}$$

Note that $\lambda_{max}(A^{-1}_{\pi_{\ell+2}}(A_{\pi_{\ell+1}}-A_{\pi_{\ell+2}})) \leq \lambda_{max}(A_{\pi_{\ell+1}}-A_{\pi_{\ell+2}})/\lambda_{min}(A_{\pi_{\ell+2}})$, there exists $C_4 > 0$ such that

$$P\{\lambda_{max}(A^{-1}_{\pi_{\ell+2}}(A_{\pi_{\ell+1}}-A_{\pi_{\ell+2}})) \geq C_4\} < C_1 N^{-(L+2)}. \tag{2.37}$$

If $\lambda_{max}(A^{-1}_{\pi_{\ell+2}}(A_{\pi_{\ell+1}}-A_{\pi_{\ell+2}})) < C_4$, there exists $C_5 > 0$ such that

$$\frac{N}{2}(\log|A_{\pi_{\ell+1}}|-\log|A_{\pi_{\ell+2}}|) = \frac{N}{2}\log|\, I + A^{-\frac{1}{2}}_{\pi_{\ell+2}}(A_{\pi_{\ell+1}}-A_{\pi_{\ell+2}})A^{-\frac{1}{2}}_{\pi_{\ell+2}}|$$

$$\geq \frac{N}{2}\log(1+\lambda_{max}[A^{-1}_{\pi_{\ell+2}}(A_{\pi_{\ell+1}}-A_{\pi_{\ell+2}})])$$

$$\geq \frac{N}{2} C_5\, \lambda_{max}[A_{\pi_{\ell+1}}-A_{\pi_{\ell+2}})/\lambda_{max}(A_{\pi_{\ell+2}}) . \tag{2.38}$$

But

$$P\{\frac{N}{2}(\log|A_{\pi_{\ell+1}}|-\log|A_{\pi_{\ell+2}}|) < C_2 D_N\}$$

$$\leq P\{\lambda_{max}[A^{-1}_{\pi_{\ell+2}}(A_{\pi_{\ell+1}}-A_{\pi_{\ell+2}})] \geq C_4\} + P\{\frac{N}{2} C_5 \frac{\lambda_{max}(A_{\pi_{\ell+1}}-A_{\pi_{\ell+2}})}{\lambda_{max}(A_{\pi_{\ell+2}})} < C_2 D_N\} \tag{2.39}$$

From (2.39), (2.37) (2.35) and (2.30), it follows that (2.31) is true.

3. ESTIMATION OF CHANGE POINTS WHEN q IS KNOWN

Let $0 < t_1 < \dots < t_q$ be the change points of the process $\underset{\sim}{x}(t)$. We can find a partition $\pi_0^{(N)} = (k_1^{(N)},\dots,k_q^{(N)})$ for each N, such that

$$E\underset{\sim}{x}(\tfrac{i}{N}) = \underset{\sim}{\mu}_j \quad \text{if } k_{j-1}^{(N)} < i \le k_j^{(N)},\ 1 \le j \le q+1, \tag{3.1}$$

where $k_0^{(N)} = 0$, $k_{q+1}^{(N)} = N$, and $\underset{\sim}{\mu}_{j-1} \ne \underset{\sim}{\mu}_j$, $1 \le j \le q+1$.

In order to simplify notation, we write $\underset{\sim}{x}_j$ for $\underset{\sim}{x}(\frac{j}{N})$, π_0 for $\pi_0^{(N)}$ and k_j for $k_j^{(N)}$.

For the integer interval [0,N], there exist $\binom{N-1}{q}$ different integer partitions denoted by K_q. Assume that $\pi = (k_1',\dots,k_q') \in K_q$, and h_π is a hypothesis such that

$$E\underset{\sim}{x}_i = \underset{\sim}{\mu}_j^* \quad \text{if } \ k_{j-1}' < i \le k_j',\ 1 \le j \le q+1,\ k_0' = 0,\ k_{q+1}' = N.$$

The model M_π is the one for which h_π is true. Let $\Theta = \{(\pi,\Sigma):\pi \in K_q, \Sigma > 0\}$ be a parametric space. We are interested in selecting one model based upon observations $\underset{\sim}{x}_1,\dots,\underset{\sim}{x}_N$. Under h_π, the logarithm of the likelihood function is

$$\log L(\theta) = -\frac{N}{2}\log|\Sigma| - \frac{1}{2}\sum_{j=1}^{q+1} \mathrm{tr}(\Sigma^{-1}A_j(N)) \tag{3.2}$$

where

$$A_j(N) = \frac{1}{N}\sum_{i=k_{j-1}'+1}^{k_j'} (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{k_{j-1}'k_j'})(\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_{k_{j-1}'k_j'})', \tag{3.3}$$

$j = 1,2,\dots,q+1$, $k_0' = 0$, $k_{q+1}' = N$, and $\theta = (\pi,\Sigma)$. Let

$$A_\pi(N) = \sum_{j=1}^{q+1} A_j(N). \tag{3.4}$$

It is not difficult to see that

$$\sup_{\theta \in \Theta_\pi} \log L(\theta) = -\frac{N}{2}\log|A_\pi(N)| - \frac{Np}{2} \tag{3.5}$$

where Θ_π denotes the parametric space under h_π

and $|A|$ denote the determinant of the matrix of the matrix A. Now let

$$G_\pi = -\frac{N}{2}\log|A_\pi(N)|, \tag{3.6}$$

where $\pi \in K_q$. Then there exists a partition $\hat{\pi} = (\hat{k}_1,\ldots,\hat{k}_q) \in K_q$ such that

$$G_{\hat{\pi}} = \max_{\pi\in K_q} G_\pi \tag{3.7}$$

Using $(\frac{\hat{k}_1}{N},\ldots,\frac{\hat{k}_q}{N})$ as an estimate of the true location of $(t_1,\ldots,t_q)$, we can prove the following theorem:

Theorem 3.1. Let $\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_N$ be a sample of size N drawn from the process $\underset{\sim}{x}(t)$ in equal space, where $\underset{\sim}{x}(t)$ is defined by (2.1) Then $(\frac{\hat{k}_1}{N},\ldots,\frac{\hat{k}_q}{N})$ is a strongly consistent estimate of $(t_1,\ldots,t_q)$.

Proof. For each N, let $\pi_o = (k_1,\ldots,k_q)$ be a partition so that (3.1) holds. It is easy to see that

$$\left|\frac{k_j}{N} - t_j\right| < \frac{1}{N},\ j = 1,2,\ldots,q \tag{3.8}$$

Now, take constants D_N which satisfy (2.28) and (2.29). Then the adjacent intervals $(k_j - D_N, k_j + D_N)$ and $(k_{j+1} - D_N,\ k_{j+1} + D_N)$ are non-intersecting for large N and all $j = 1,2,\ldots,q$. Define $\tilde{K}_q^{(N)} = \{\pi' = (k_1',\ldots,k_q'):\exists\, j \le q$ such that $\{k_1',\ldots,k_q'\} \cap \{k_j-D_N,\ k_j+D_N\} = \phi\}$. We shall prove that with probability one for large N,

$$G_{\pi'} < G_{\pi_o} \text{ for all } \pi' \in \tilde{K}_q^{(N)}.$$

Take a partition $\pi' = (k_1',\ldots,k_q') \in \tilde{K}_q^{(N)}$, we can construct a new partition π_1 which has all cut-points of both π_o and π' except the point k_j. Denote π_1 by $(k_1'',\ldots,k_\ell'')$. Assume k_j drops into (k_i'',k_{i+1}''). Let $\pi_2 = (k_1'',\ldots,k_i'',k_j,k_{i+1}'',\ldots,k_\ell'')$. From Lemma 3.4, with probability for large N we have

$$P\{G_{\pi_2} - G_{\pi_1} \le C_2 D_N\} < C_1 N^{-(L+2)}. \tag{3.9}$$

Since $\pi_1 = (k_1'',\ldots,k_\ell'')$ is a refinement partition of π', it is easily proved that

$(A_{\pi'} - A_{\pi_1})$ is a non-negative definite matrix based on the expressions of A_{π_1} and A_{π}'. So

$$G_{\pi'} \leq G_{\pi_1}. \tag{3.10}$$

At the same time, for $\pi_2 \prec \pi_o$, we can sequentially construct $\pi_2 = (k''_1, \ldots, k''_i, k_j, k''_{i+1}, \ldots, k'')$ from $(k_1, \ldots, k_q) = \pi_o$ by means of adding one cut-point of π' each time. Let π_{q+1} denote a partition which has all cut-points of π_o and one cut-point of π' different from $k_1, \ldots, k_q$. π_{q+2} a partition which has all cut-points of π_{q+1} and one cut-point of π' different from all cut-points of $\pi_{q+1}, \ldots$. Thus we obtain

$$\pi_o = \pi_q \succ \pi_{q+1} \succ \pi_{q+2} \succ \cdots \succ \pi_{\ell+1} = \pi_2.$$

By Lemma 2.3

$$P\{G_{\pi_{q+j+1}} - G_{\pi_{q+j}} \geq C \log N\} < C_1 N^{-(L+2)}, \text{ for } j = 0,1,\ldots,\ell-q, \quad \text{a.s.}$$

Hence $\pi_o \succ \pi_2$ implies

$$P\{G_{\pi_2} - G_{\pi_o} \geq LC\log N\} < LC_1 N^{-(L+2)}. \tag{3.11}$$

Combining (3.9) - (3.11) and noticing

$$G_{\pi'} - G_{\pi_0} = (G_{\pi'} - G_{\pi_1}) + (G_{\pi_1} - G_{\pi_2}) + (G_{\pi_2} - G_{\pi_0}),$$

we have for large N

$$P\{G_{\pi'} - G_{\pi_o} \geq - \frac{C_2}{2} D_N\} < (L+1)C_1 N^{-(L+2)}. \tag{3.12}$$

Note that the constatns C_1 and C_2 are independent of the choice of $\pi' \varepsilon \tilde{K}_q^{(N)}$. Since $\#(\tilde{K}_q^{(N)}) \leq N^L$, we have

$$P\{G_{\pi'} - G_{\pi_o} \geq 0 \text{ for at least one } \pi' \varepsilon \tilde{K}_q^{(N)}\} < C_1 (q+1) N^{-2}.$$

By Borel-Cantelli's lemma, with probability one for large N,

$$\hat{\pi} = (\hat{k}_1, \ldots, \hat{k}_q) \notin \tilde{K}_q^{(N)},$$

which implies that

$$\left|\frac{\hat{k}_j}{N} - t_j\right| < \frac{D_N + 1}{N},$$

Thus the theorem is proved.

4. ESTIMATES OF THE NUMBER AND LOCATIONS OF CHANGE POINTS WHEN q IS UNKNOWN

When the number of the change points is unknown, we have to estimate q. For this purpose, some notations are introduced. Suppose that $\pi_\ell^{(N)} = (k_1^{(N)},\dots,k_\ell^{(N)})$ is an integer partition of [0,N]. Set

$$K = \{(k_1^{(N)},\dots,k_\ell^{(N)}),\ 0 \le \ell \le L,\ 0 < k_1^{(N)} < \dots < k_\ell^{(N)} < N\} \tag{4.1}$$

where $\ell \le L$ and L is a constant. $\pi_\ell^{(N)}$, $k_j^{(N)}$ will be written as π_ℓ, k_j respectively below.

Let the sequences C_N and D_N satisfy the following conditions:

$$\lim_{N\to\infty} \frac{C_N}{\log N} = \infty,\quad \lim_{N\to\infty} \frac{D_N}{C_N} = \infty,\quad \lim_{N\to\infty} \frac{D_N}{N} = 0. \tag{4.2}$$

Set

$$H_{\pi_\ell} = -\frac{N}{2}\log |A_{\pi_\ell}(N)| - \ell C_N. \tag{4.3}$$

Assume $\hat{\pi}^{(N)} = (\hat{k}_1^{(N)},\dots,\hat{k}_{\hat{h}^{(N)}}^{(N)}) \in K$ is a partition of maximizing H_{π_ℓ} when ℓ runs over all integers of [0,L], i.e.,

$$H(\hat{k}_1^{(N)},\dots,\hat{k}_{\hat{h}^{(N)}}^{(N)}) = \max_{\pi_\ell \in K} H_{\pi_\ell}(N). \tag{4.4}$$

Then, $\hat{k}_1^{(N)},\dots,\hat{k}_{\hat{h}^{(N)}}^{(N)}$ are grouped into some groups by the following procedure: Let $\hat{k}_1^{(N)}$ belong to the first group, say $M_1^{(N)}$. If $\hat{k}_2^{(N)} - \hat{k}_1^{(N)} < D_N$, $\hat{k}_2^{(N)}$ also belongs to $M_1^{(N)}$; otherwise $\hat{k}_2^{(N)}$ is called an element of the second group $M_2^{(N)}, \dots$. In general, assume $\hat{k}_\ell^{(N)} \in M_i$. Then

$$\hat{k}_{\ell+1}^{(N)} \in \begin{cases} M_i^{(N)}, & \text{if } \hat{k}_{\ell+1}^{(N)} - \hat{k}_\ell^{(N)} < D_N \\ M_{i+1}^{(N)}, & \text{otherwise.} \end{cases} \tag{4.5}$$

Based on this procedure, we get groups $M_1^{(N)},\dots,M_{\hat{q}_N}^{(N)}$ finally. Let $\hat{k}_{i_j}^{(N)}$ be an element of $M_j^{(N)}$, $j = 1,2,\dots,\hat{q}_N$. Then we have the following theorem:

<u>Theorem 4.1</u>. Under h_{π_o}, we have

$$\text{(i)} \quad \hat{q}_N \to q, \quad \text{a.s.} \tag{4.6}$$

$$\text{(ii)} \quad \left(\frac{\hat{k}^{(N)}_{i_1}}{N}, \ldots \frac{\hat{k}^{(N)}_{i_{\hat{q}(N)}}}{N}\right) \to (t_1, \ldots, t_q), \quad \text{a.s.} \tag{4.7}$$

Proof. In order to simplify notation, we write $\hat{k}_j$ for $\hat{k}_j^{(N)}$, π_ℓ for $\pi_\ell^{(N)}$, etc. If the hypothesis h_{π_o} is true, then there exists a partition $\pi_o = (k_1, \ldots, k_q)$ so that (3.1) holds. From (3.11), for some $C > 0$,

$$\sum_{\pi' \in K, \pi' \prec \pi_o} P\{G_{\pi'_\ell} - G_{\pi_o} \geq CL\log N\} < \infty$$

Then, with probability one for large N, we have for any $\ell \leq L$ and $\pi'_\ell = (j_1, \ldots, j_\ell)\ \pi$

$$H_{\pi'_\ell} - H_{\pi_o} = G_{\pi'_\ell} - G_{\pi_o} - \ell C_N + qC_N = O(\log N) - (\ell - q)C_N$$

$$< -\frac{1}{2} C_N < 0. \tag{4.8}$$

This fact shows that the maximum of H_π cannot reach at refinement partitions of π_o. Next, set $R_N = D_N^{\frac{1}{4}} C_N^{\frac{3}{4}}$. It follows that

$$\lim_{N\to\infty} \frac{R_N}{D_N} = 0, \quad \lim_{N\to\infty} \frac{C_N}{R_N} = 0, \text{ and } \lim_{N\to\infty} \frac{R_N^2}{D_N C_N} = 0. \tag{4.9}$$

Define

$$\tilde{K}^{(N)} = \{\pi' = (k'_1, \ldots, k'_\ell) \in K : \exists\, j \leq q \text{ such that } |k'_i - k_j| > R_N \text{ for } i = 1, \ldots, \ell\}$$

Then, following the same lines of obtaining (3.12), we have, with probability one for large N,

$$\sum_{\pi' \in \tilde{K}^{(N)}} P\{G_{\pi'} - G_{\pi_o} \geq -\frac{C_2}{4} R_N\} < \infty \tag{4.10}$$

and with probability one for large N, we have for any $\pi'_\ell = (k'_1,\ldots,k'_\ell) \in \tilde{K}^{(N)}$,

$$H_{\pi'_\ell} - H_{\pi_o} = G_{\pi'_\ell} - G_{\pi_o} - (\ell-q)C_N < -\frac{C_2}{4} R_N + |\ell-q|C_N < -\frac{C_2}{8} R_N < 0. \qquad (4.11)$$

This inequality also shows that the maximum of H_π cannot reach at π' above. In other words, if $\hat{\pi} = (\hat{k}_1,\ldots,\hat{k}_{\hat{h}})$ is a partition of these maximizing H_π, $\pi \in K$, then at least one cut-point of $\hat{\pi}$ falls into each R_N-neighbor of k_j, $1 \le j \le q$, with probability one for large N. Hence, based on our procedure of grouping $\hat{k}_1,\ldots,\hat{k}_{\hat{h}}$, it is hnown that, with probability one for large N,

$$\hat{q}_N \ge q. \qquad (4.12)$$

Further, we prove that all cut-points of $\hat{\pi} = (\hat{k}_1,\ldots,\hat{k}_{\hat{h}})$ must drop into $\bigcup_{j=1}^{q} (k_j - \frac{D_N}{2}, k_j + \frac{D_N}{2})$, with probability one for large N.

First, suppose $\pi'_\ell = (k'_1,\ldots,k'_\ell)$ satisfies the following condition:

$$\pi'_\ell \in K - \tilde{K}^{(N)}, \; \pi_\ell \quad \pi_o \text{ and for some } i \le \ell,$$

$$k'_b \in \bigcup_{m=1}^{q} (k_m - R_N, k_m + R_N), \quad b = 1,2,\ldots,i-1,i+1,\ldots,\ell \qquad (*)$$

and

$$k'_i \notin \bigcup_{m=1}^{q} (k_m - R_N/2, k_m + R_N/2).$$

Write $\pi''_{\ell-1} = (k'_1,\ldots,k'_{i-1},k'_{i+1},\ldots,k'_\ell)$. It is obvious that $\pi'_\ell \quad \pi''_{\ell-1}$. Since

$$A_{\pi''_{\ell-1}} - A_{\pi'_\ell} = \frac{1}{N}(k'_i - k'_{i-1})(\bar{x}_{k'_{i-1}k'_i} - \bar{x}_{k'_{i-1}k'_{i+1}})(\bar{x}_{k'_{i-1}k'_i} - \bar{x}_{k'_{i-1}k'_{i+1}})'$$

$$+ \frac{1}{N}(k'_{i+1} - k'_i)(\bar{x}_{k'_i k'_{i+1}} - \bar{x}_{k'_{i-1}k'_{i+1}})(\bar{x}_{k'_i k'_{i+1}} - \bar{x}_{k'_{i-1}k'_{i+1}})' \ge 0 \qquad (4.13)$$

Assume that $k'_{i-1} < k_{j-1} < k'_i < k'_{i+1} < k_j$ (see figure 1)(For other permutations of $k'_{i-1}, k'_i, k'_{i+1}, k_{j-1}$ and k_j, the treatment is similar.)

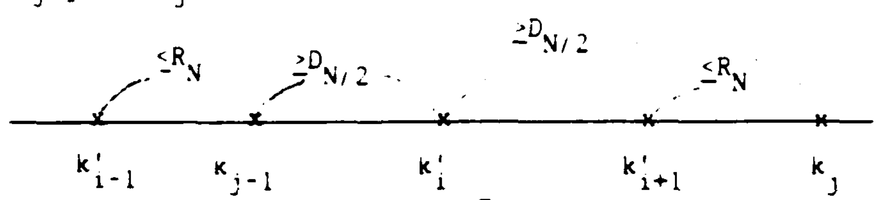


Figure 1

Set

$$\alpha = \frac{k_{j-1} - k'_{i-1}}{k'_i - k'_{i-1}}, \quad \beta = 1 - \alpha.$$

Making routine calculation and noticing (2.14), we can change (4.13) into the form.

$$A_{\pi''_{\ell-1}} - A_{\pi'_\ell} = \frac{(k'_i - k'_{i-1})(k'_{i+1} - k'_i)}{N(k'_{i+1} - k'_{i-1})} (\alpha \bar{x}_{k'_{i-1} k_{j-1}} + \beta \bar{x}_{k_{j-1} k_j} - \bar{x}_{k'_i k'_{i+1}}).$$

$$\cdot (\alpha \bar{x}_{k'_{i-1} k_{j-1}} + \beta \bar{x}_{k_{j-1} k'_i} - \bar{x}_{k'_i k'_{i+1}})'$$

$$\le \frac{2(k'_i - k'_{i-1})(k'_{i+1} - k'_i)}{N(k'_{i+1} - k'_{i-1})} \alpha^2 (\mu_{j-1} - \mu_j)(\mu_{j-1} - \mu_j)'$$

$$+ \frac{16(k'_i - k'_{i-1})(k'_{i+1} - k'_i)}{N(k'_{i+1} - k'_{i-1})} [\alpha^2 (\bar{x}_{k'_{i-1} k_{j-1}} - \mu_{j-1})(\bar{x}_{k'_{i-1} k_{j-1}} - \mu_{j-1})'$$

$$+ \beta^2 (\bar{x}_{k_{j-1} k'_i} - \mu_j)(\bar{x}_{k_{j-1} k'_i} - \mu_j)' + (\bar{x}_{k'_i k'_{i+1}} - \mu_j)(\bar{x}_{k'_i k'_{i+1}} - \mu_j)'$$

$$\hat{=} I_1 + I_2 \tag{4.14}$$

But by Lemma 4.3, we get uniformly for all those π'_ℓ satisfying (*),

$$\lambda_{max}(I_2) = 0(\frac{\log N}{N}) = o(\frac{C_N}{N}) \quad \text{a.s.} \tag{4.15}$$

and

$$\lambda_{max}(I_1) = 0(\frac{(k_{j-1} - k'_{i-1})^2 (k'_{i+1} - k'_i)}{N(k'_i - k'_{i-1})(k'_{i+1} - k'_{i-1})}) \tag{4.16}$$

Since $k_{j-1} - k'_{i-1} < R_N$, $k'_i - k_{j-1} > \frac{D_N}{2}$, from (4.2), for large N, $k'_i - k'_{i-1} = (k'_i - k_{j-1}) + (k_{j-1} - k'_{i-1}) > \frac{D_N}{2}$, we get uniformly for all those π'_ℓ satisfying (*),

$$\lambda_{max}(I_1) = 0(\frac{R_N^2}{ND_N}) = o(\frac{C_N}{N}), \quad \text{a.s.} \tag{4.17}$$

and

$$\lambda_{\max}(A_{\pi''_{\ell-1}} - A_{\pi'_\ell}) \le \lambda_{\max}(I_1) + \lambda_{\max}(I_2) = o(\frac{C_N}{N}), \quad \text{a.s.} \tag{4.18}$$

From this inequality and Remark 2.2, with probability one for large N, we have for all those π'_ℓ satisfying (*),

$$H_{\pi''_{\ell-1}} - H_{\pi'_\ell} = -\frac{N}{2}(\log|A_{\pi''_{\ell-1}}| - \log|A_{\pi'_\ell}|) - (\ell-1)C_N + \ell C_N$$

$$= o(C_N) + C_N > \frac{1}{2}C_N > 0 \tag{4.19}$$

Next, set $\pi'_\ell \in K - \tilde{K}^{(N)}$. Assume that $\pi'_\ell \not\succ \pi_o$ and there exist more than one cut-points of π'_ℓ in the

complementary set of $\bigcup_{j=1}^{q} (k_j - \frac{D_N}{2}, k_j + \frac{D_N}{2})$. Write

$$\pi'_\ell = (k'_1, \ldots, k'_i, k'_{i+1}, \ldots, k'_{i+\omega+1}, \ldots, k'_\ell)$$

and suppose that there exists a cut-point, k_j say, of π_o such that

$$k'_i \in (k_{j-1} - \frac{D_N}{2}, k_{j-1} + \frac{D_N}{2})$$

$$k'_{i+\omega+1} \in (k_j - \frac{D_N}{2}, k_j + \frac{D_N}{2})$$

and

$$k'_{i+1}, \ldots, k'_{i+\omega} \in [k_{j-1} + \frac{D_N}{2}, k_j - \frac{D_N}{2}]$$

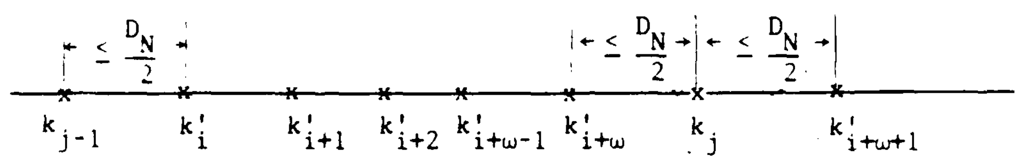


Figure 2

Let

$$\pi'_{\ell-\omega+m+1} = (k'_1, \ldots, k'_i, k'_{i+1}, \ldots, k'_{i+m}, k'_{i+\omega}, k'_{i+\omega+1}, \ldots, k'_\ell) \quad m = 0, 1, \ldots, \omega-1.$$

$$\pi'_{\ell-\omega} = (k'_1,\ldots,k'_i,\ k'_{i+\omega+1},\ldots,k'_\ell).$$

From (4.18), we get uniformly for all those π'_ℓ and $\pi'_{\ell-\omega+m}$ above,

$$\lambda_{max}(A_{\pi'_{\ell-\omega}} - A_{\pi'_{\ell-\omega+1}}) = o(\frac{C_N}{N}),\ \text{a.s.}$$

Following the same lines as obtaining (4.15), we have uniformly for all those $\pi'_{\ell-\omega+m}$ above,

$$\lambda_{max}(A_{\pi'_{\ell-\omega+2}} - A_{\pi'_{\ell-\omega+1}}) = o(\frac{C_N}{N}),\quad \text{a.s.}$$

Again notice that $Ex_{k'_r} = \mu_j$, $r = i+1,\ldots,i+\omega$, hence by Lemma 2.3, we have uniformily for $\pi'_{\ell-\omega}$ and π'_ℓ,

$$\lambda_{max}(A_{\pi'_{\ell-\omega}} - A_{\pi'_\ell}) \leq \sum_{m=0}^{\omega-1} \lambda_{max}(A_{\pi'_{\ell-\omega+m}} - A_{\pi'_{\ell-\omega+m+1}})$$

$$\leq o(\frac{C_N}{N}) + (\omega-2)\ O(\frac{\log N}{N}) = o(\frac{C_N}{N}),\quad \text{a.s.}$$

From this fact and the Remark 2.2 of Lemma 2.3, it is proved that with probability one for large N, we have for all those π'_ℓ above, there exists $\pi_{\ell-\omega}$ such that

$$H_{\pi'_{\ell-\omega}} - H_{\pi'_\ell} = -\frac{N}{2}(\log|A_{\pi'_{\ell-\omega}}| - \log|A_{\pi'_\ell}|) - (\ell-h)C_N + \ell C_N$$

$$= o(C_N) + hC_N > \frac{1}{2} C_N > 0. \tag{4.20}$$

The inequalities (4.19) and (4.20) show that the maximum of H_π, $\pi \in K$, can not reach at partitions for which there exists one cut-point at least droping out of $\bigcup_{j=1}^{q} (k_j - \frac{D_N}{2}, k_j + \frac{D_N}{2})$. According to our procedure of grouping $\hat{k}_1,\ldots,\hat{k}_{\hat{h}}$, with probability one for large N, no two or more groups exist in each $\frac{D_N}{2}$-neighbor of k_j. Hence

$$\hat{q}_N \leq q \quad \text{a.s.} \tag{4.21}$$

Combining (4.12) and (4.21) we have, with probability one for large N,

$$\hat{q}_N = q.$$

It is easy to see that

$$\frac{\hat{k}_{i_j}}{N} \to t_j \quad \text{a.s.}$$

From that $\lim_{N\to\infty} \frac{D_N}{N} = 0$. Thus the theorem is proved.

5. TESTS OF HYPOTHESES FOR DETECTION OF CHANGE POINTS

First, we assume that Σ is known and the number of change points is unknown. Let H_i: $\underset{\sim}{\mu}_i = \underset{\sim}{\mu}_{i+1}$ for i = 1, 2,..., N-1. Then we can test (see Krishnaiah (1969)) the hypotheses $H_1, \ldots, H_{N-1}$ simultaneously as follows. Let

$$\chi_i^2 = (\underset{\sim}{x}_i - \underset{\sim}{x}_{i+1})'(2\Sigma)^{-1}(\underset{\sim}{x}_i - \underset{\sim}{x}_{i+1})$$

for $i = 1,2,\ldots,(N-1)$. Then, we accept or reject H_i according as

$$\chi_i^2 \lessgtr c_\alpha$$

where

$$P[\chi_i^2 \leq c_\alpha;\ i=1,2,\ldots,N-1 \mid \bigcap_{i=1}^{N-1} H_i] = (1-\alpha).$$

The joint distribution of $\chi_1^2, \ldots, \chi_{N-1}^2$ is a multivariate chi-square distribution. Tables for approximate percentage points of c_α are given in Krishnaiah (1980). If Σ is unknown but an independent estimate (S/ν) of Σ is available, we use T_i^2 instead of χ_i^2 as test statistics where

$$T_i^2 = (\underset{\sim}{x}_i - \underset{\sim}{x}_{i+1})'(S/\nu)^{-1}(\underset{\sim}{x}_i - \underset{\sim}{x}_{i+1}).$$

We can determine the number of change points and estimate the locations of change points by the above method. Here, we note that $\bigcap_{i=1}^{N-1} H_i$ indicates that $\underset{\sim}{\mu}_1 = \ldots = \underset{\sim}{\mu}_N$ and no change points exist. If $H_1, \ldots, H_{N-1}$ are all simultaneously rejected, then we have (N-1) change points. If q of the hypothesis $H_i (i=1,\ldots, N-1)$ are rejected, then there are q change points. Suppose $\bigcap_{j=1}^{i} H_j$ is accepted and H_{i+1} is rejected, a change point occurs at t_i. Suppose q is known, then $t_{\hat{k}_1}, \ldots, t_{\hat{k}_q}$ are change points where they are chosen as follows:

$$\chi^2_{\hat{k}_1} = \max\{\chi^2_1, \ldots, \chi^2_{N-1}\}$$

$$\chi^2_{\hat{k}_2} = \max\{\chi^2_i, \; i=1,\ldots,N-1; \; i \neq \hat{k}_1\}$$

$$\chi^2_{\hat{k}_3} = \max\{\chi^2_i; \; i=1,\ldots,N-1; \; i \neq \hat{k}_1, \hat{k}_2\}$$

$$\vdots \qquad \vdots$$

$$\chi^2_{\hat{k}_q} = \max\{\chi^2_i; \; i=1,\ldots,N-1; \; i \neq \hat{k}_1, \ldots, \hat{k}_{q-1}\}.$$

We can use finite intersection tests proposed by Krishnaiah (1965) for multiple comparisons of mean vectors also to estimate the locations of change points.

REFERENCES

1. Bhattacharyya, G.K. and Johnson, R.A. (1968). Nonparameteric tests for shift at an unknown time point. Ann. Math. Statist., 39, 1731-1743.

2. Booth, N.B. and Smith, A.F.M. (1982). A Bayesian approach to retrospective identification of change points. Journal of Econometrics, 19, 7-22.

3. Chernoff, H and Zacks, S. (1964). Estimating the current mean of a normal distribution which is subjected to changes in time. Ann. Math. Statist., 35, 999-1018.

4. Fuk, D.K.H. and Nagaev, S.V. (1971). Probability inequalities for sums of independent random variables. Theory of Probability and its Applications, 16, 643-660.

5. Hinkley, D.V. (1970). Inference about the change-point in a sequence of random variables. Biometrika, 57, 1-17.

6. Holbert, D. and Broemeling, L. (1977). Bayesian inferences related to shifting sequences and two-phase regression. Communications in Statistics - Theory and Methods, A6, 265-275.

7. Hudson, D.J. (1966). Fitting segmented curves whose joint points have to be estimated. J. Amer. Statist. Assoc., 61, 1097-1129.

8. Krishnaiah, P.R. (1965). Multiple comparison tests in multiresponse experiments. Sankhyā, Ser. A., 27, 35-53.

9. Krishnaiah, P.R. (1969). Some developments on simultaneous test procedures. In Developments in Statistics, Vol. 2, (P.R. Krishnaiah, editor) Academic Press, Inc. pp 157-201.

10. Krishnaiah, P.R. (1980). Computations of some multivariate distributions. In Handbook of Statistics, Vol. 1: Analysis of Variance, pp. 745-971. North-Holland Publishing Company.

11. Mazumdar, M., Sinha, B.K. and Li, C.C. (1985). A comparison of several estimates of edge point in noisy digital data across a step edge. IEEE Proceedings on Computer Vision and Pattern Recognition held at San Francisco, California, June 19-23, 1985.

12. Mazumdar, M. and Chao, R.J. (1986). Use of maximum likelihood estimate in detection of edges in a noisy digital image. To appear in Proceeding of the International Conference on Advances in Pattern Recognition and Digital Techniques held at the Indian Statistical Institute, Calcutta, January 1986.

13. Page, E.S. (1955). A test for a change in a parameter occurring at an unknown point, Biometrika. 42, 523-527.

14. Page, E.S. (1957). On problems in which a change in parameter occurs at an unknown point. Biometrika, 44, 248-252.

15. Sen, A.K. and Srivastava, M.S. (1973). On multivariate tests for detecting change in mean. Sankhyā, Ser. A, 35, 173-185.

16. Siotani, M. (1959). The extreme value of the generalized distances of the individual points in multivariate normal sample. Annals of the Institute of Statistical Mathematics, 10, 183-208.

17. Srivastava, M.S. and Worsley, K.J. (1986). Likelihood ratio tests for a change in the multivariate normal mean. Journal of the American Statistical Association, **81**, 199-204.

18. Tabatabai, A.J. and Mitchell, O.R. (1984). Edge location to subpixel values in digital imagery. IEEE Transactions on Pattern Recognition and Machine Intelligence, **6**, 188-201.

19. Vostrikova, L.J. (1981). Detecting disorder in multidimensional random processes. Soviet Mathematics Doklady, **24**, 54-59.

END

12-87

DTIC